TDECQ Calculation According to D3.1

Norman L. Swenson, Ph.D.

Affiliations: Nokia and Point2

Version 1.0

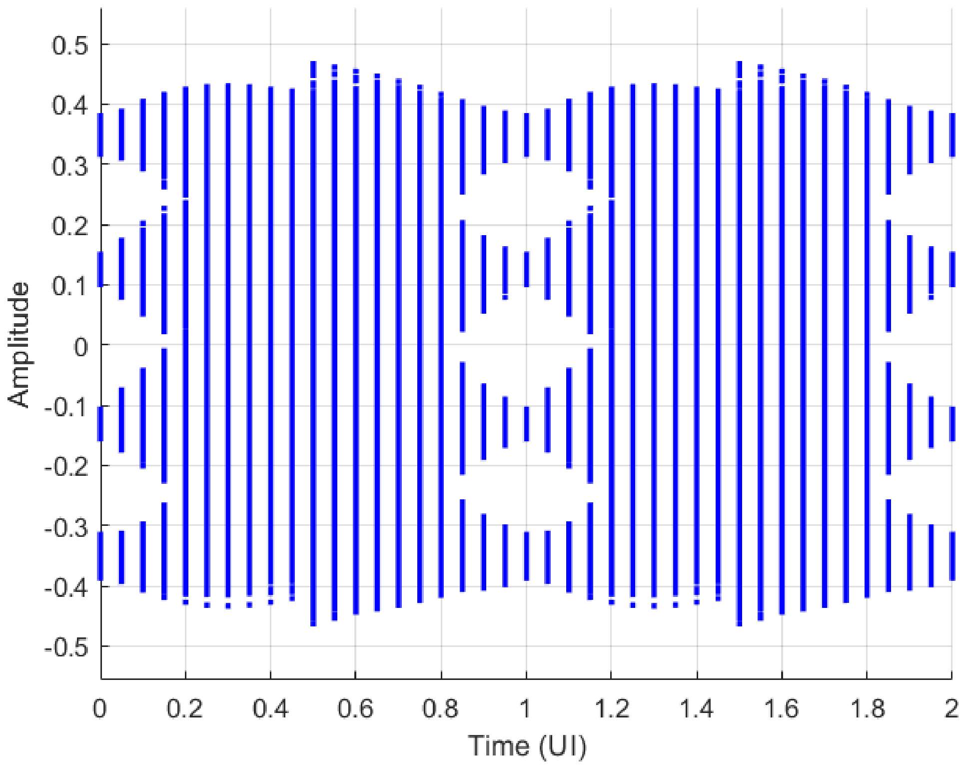
IEEE 802.3dj, IEEE 802 July Plenary Meeting

7 July 2026

Comments addressed: 95, 96, 97, 103, 205, 227, 229, 230, 263, 264, 267, 287

Eye Diagram Based on Scatter Plots

- Based on phase steps of T/20
- Density can be increased by choosing denser phase grid
- Histogram locations need not be optimal, but must be located around the discontinuity plus T/2.



Comment(s) addressed: 96, 97, 263,

Symbol Error Rate Estimation

- Refer to swenson_3dj_adhoc_01a_260624.pdf at https://www.ieee802.org/3/dj/public/adhoc/optics/0626_OPTX/swenson_3dj_adhoc_01a_260624.pdf
- Q-function-based method proposed in comments listed are straightforward, fast, easier to understand, and more accurate than the current histogram-based methodology

Comment(s) addressed: 103, 229, 230,

Methodology

- The Appendix gives detail on an MMSE algorithm that achieves repeatable results and has been implemented by several parties
- There is no evidence that blind search gives substantially better or different results.
- The qualifier that better results can be used should be eliminated.

Comment(s) addressed: 95, 205, 287

Appendix

MMSE Calculation of Reference Equalizer Taps for TDECQ

1 Introduction

This document outlines the Minimum Mean-Squared Error (MMSE) formulation for determining the reference equalizer coefficients for Clause 180.9.6 TDECQ calculations. By employing a direct Wiener-Hopf matrix solution, this approach eliminates the need for computationally intensive brute-force equalizer coefficient searches.

1.1 Alphabet Definition and Scope Limitation

- **Sequence Alphabet Definition:** The ideal transmitted SSPRQ target sequence $x(n)$ is mapped utilizing the standard symmetric, zero-mean PAM4 integer alphabet $\{-3, -1, 1, 3\}$.
- **Scope Limitation regarding Tap Bounds:** This mathematical derivation represents an unconstrained global optimization baseline. The strict tap coefficient boundary limits outlined in Table 180-16 are not enforced in this derivation. If the unconstrained solution yields coefficients exceeding the tap limits, the final tap values must instead be determined using a bounded, constrained optimization framework such as quadratic programming.

The reference receiver architecture uses a Decision-Feedback Equalizer (DFE) with 15 feedforward taps and 1 feedback tap. Assuming correct decisions are fed back under a known test pattern sequence, the feedback path uses the ideal transmitted symbols. The captured waveform samples natively incorporate deterministic inter-symbol interference (ISI) and signal-dependent transmitter noise, while background Additive White Gaussian Noise (AWGN) is shaped by a 4-pole Bessel-Thomson filter.

2 System Model and Notation

Following the notation of IEEE Draft P802.3dj/D3.1, the total continuous-time input signal to the sampler is expressed as:

$$z(t) = r(t) + \eta(t) \tag{1}$$

where $r(t)$ is the received optical signal captured by the sampling scope—which has already undergone filtering by the 4-pole Bessel-Thomson filter and naturally includes any transmitter noise—and $\eta(t)$ represents the filtered background Gaussian noise.

The discrete-time signal sampled at the symbol rate with a sampling phase ϕ_0 is given by:

$$z_n = z(nT + \phi_0) \tag{2}$$

The reference equalizer computes its output y_n using 15 feedforward taps $w(a+i)$ for $i \in \{0, 1, \ldots, 14\}$ and a single feedback tap $b(1)$:

$$y_n = \sum_{i=0}^{14} w(a+i) z_{n-a-i} - b(1)x(n-1) \tag{3}$$

where $x(n)$ denotes the ideal transmitted SSPRQ symbol sequence, and $a \in \{-3, -2, -1, 0\}$ denotes the number of pre-cursor taps.

We establish a compact matrix formulation by defining the compound input vector $\mathbf{z}_n$ and the consolidated coefficient vector $\mathbf{w}$:

$$\mathbf{z}_n = \begin{bmatrix} z_n & z_{n-1} & \ldots & z_{n-14} & x(n-1) \end{bmatrix}^T \tag{4}$$

$$\mathbf{w} = \begin{bmatrix} w(a) & w(a+1) & \ldots & w(a+14) & -b(1) \end{bmatrix}^T \tag{5}$$

This allows the equalizer output expression to be written as a vector inner product:

$$y_n = \mathbf{w}^T \mathbf{z}_n \tag{6}$$

3 MMSE Formulation and Wiener-Hopf Solution

The error signal e_n relative to the aligned target symbol $x(n)$ is defined as:

$$e_n = y_n - x(n) = \mathbf{w}^T \mathbf{z}_n - x(n) \tag{7}$$

$$J(\mathbf{w}) = E\left[e_n^2\right] = E\left[\left(\mathbf{w}^T \mathbf{z}_n - x(n)\right)^2\right] \tag{8}$$

Expanding the expectation yields:

$$J(\mathbf{w}) = \mathbf{w}^T E\left[\mathbf{z}_n \mathbf{z}_n^T\right] \mathbf{w} - 2\mathbf{w}^T E\left[\mathbf{z}_n x(n)\right] + E\left[x(n)^2\right] \tag{9}$$

Defining the input autocorrelation matrix $\mathbf{R}_{zz} = E\left[\mathbf{z}_n \mathbf{z}_n^T\right]$ and the cross-correlation vector $\mathbf{p}_{zx} = E\left[\mathbf{z}_n x(n)\right]$, the objective function simplifies to:

$$J(\mathbf{w}) = \mathbf{w}^T \mathbf{R}_{zz} \mathbf{w} - 2\mathbf{w}^T \mathbf{p}_{zx} + E\left[x(n)^2\right] \tag{10}$$

Taking the gradient with respect to the tap vector $\mathbf{w}$ and setting it to zero minimizes the cost function:

$$\nabla_{\mathbf{w}} J(\mathbf{w}) = 2\mathbf{R}_{zz}\mathbf{w} - 2\mathbf{p}_{zx} = \mathbf{0} \tag{11}$$

This results in the classic Wiener-Hopf matrix equation:

$$\mathbf{R}_{zz}\mathbf{w}_{\text{opt}} = \mathbf{p}_{zx} \implies \mathbf{w}_{\text{opt}} = \mathbf{R}{zz}^{-1}\mathbf{p}_{zx} \tag{12}$$

4 Partitioned Matrix Structure and Sequence Constraints

The global (16×16) correlation matrix $\mathbf{R}_{zz}$ partitions into sub-blocks isolating the feedforward and feedback dimensions:

$$\mathbf{R}_{zz} = \left[\begin{array}{c|c} \mathbf{R}_{zz,\mathrm{f}} & \mathbf{R}_{zx,\mathrm{fb}} \\ \hline \mathbf{R}_{xz,\mathrm{fb}} & R_{xx} \end{array}\right] \tag{13}$$

4.1 Composition of the Feedforward Sub-block $\mathbf{R}_{zz,\mathrm{f}}$

The 15×15 sub-matrix $\mathbf{R}_{zz,\mathrm{f}}$ captures the combined statistics of the captured signal and background noise. Because $r(t)$ and $\eta(t)$ are mutually uncorrelated, the matrix splits into:

$$\mathbf{R}_{zz,\mathrm{f}} = \mathbf{R}_{rr} + \mathbf{R}_{\eta\eta} \tag{14}$$

4.1.1 Computation of $\mathbf{R}_{rr}$ from Captured Samples

Because $r(t)$ corresponds to the actual captured noisy waveform samples from the oscilloscope, it inherently includes any transmitter noise contributions alongside the deterministic ISI profile. For a periodic SSPRQ test pattern of length $N = 65,535$, the indices $j, k \in \{0, 1, \ldots, 14\}$ of $\mathbf{R}_{rr}$ are evaluated directly via time-averaging:

$$[\mathbf{R}_{rr}]_{j,k} = \frac{1}{N} \sum_{n=1}^{N} r(nT - jT + \phi_0) r(nT - kT + \phi_0) \tag{15}$$

4.1.2 Derivation and Numerical Baseline of Background Noise Matrix $\mathbf{R}_{\eta\eta}$

The background noise matrix $\mathbf{R}_{\eta\eta}$ is a symmetric Toeplitz matrix uniquely determined by its first row vector $\mathbf{r}_{\eta\eta} = [R_{\eta\eta}(0), R_{\eta\eta}(1), \ldots, R_{\eta\eta}(14)]$. The continuous-time transfer function of a fourth-order low-pass Bessel-Thomson filter is defined by:

$$H_{\mathrm{BT}}(s) = \frac{105}{s^4 + 10s^3 + 45s^2 + 105s + 105} \tag{16}$$

To evaluate the normalization matching the specified 53.125 GHz −3 dB electrical power frequency cutoff at a symbol rate of 106.25 GBaud, the frequency scale factor maps via $s = j2\pi f \cdot \tau_0$ with $\tau_0 = \frac{2.113915}{2\pi \cdot 53.125 \times 10^9}$.

The continuous power spectral density of the filtered Gaussian process is proportional to $|H_{\mathrm{BT}}(f)|^2$. Applying the Wiener-Khinchin theorem, the autocorrelation function at a continuous time lag τ is found by taking the inverse Fourier transform:

$$R_{\eta\eta}(\tau) = \sigma_G^2 \cdot \frac{\int_{\infty}^{\infty} |H_{\mathrm{BT}}(f)|^2 e^{j2\pi f\tau} df}{\int_{-\infty}^{\infty} |H_{\mathrm{BT}}(f)|^2 df} \tag{17}$$

Sampling this correlation profile at uniform discrete symbol intervals $\tau = mT$, where $T = \frac{1}{106.25 \text{ GBaud}} \approx 9.41176$ ps, isolates the integer lag values $m = |j - k|$. Setting the total background noise power on-diagonal to $R_{\eta\eta}(0) = \sigma_G^2$, numerical integration of the normalized frequency domain structure yields the following unique row coefficients:

$$\mathbf{r}_{\eta\eta} = \sigma_G^2 \left[\; 1.0000, \quad 0.0206, \quad 0.0013, \quad 0.0151, \quad -0.0001, \quad \mathbf{0}_{1\times 10} \;\right] \tag{18}$$

where any entry for element index lag $m \geq 5$ decays completely below 10^{-4} and is truncated to zero.

4.2 Deterministic SSPRQ Cross-Correlation and Feedback Elements

Because SSPRQ is not perfectly independent and identically distributed (i.i.d.), the remaining variables must be computed directly from the pattern's structural indices:

- The true sequence variance is computed via:

$$R_{xx} = \frac{1}{N}\sum_{n=1}^{N} x^2(n-1) \tag{19}$$

- The j-th entry of the 15×1 cross-correlation block $\mathbf{R}_{zx,\mathrm{fb}}$ captures the non-zero correlation between the past transmitted symbol and the captured filtered waveform:

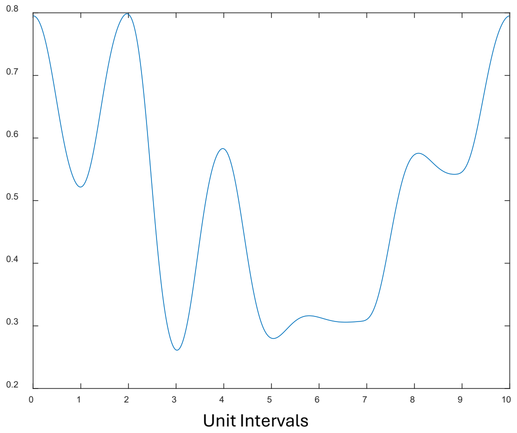
$$[\mathbf{R}_{zx,\mathrm{fb}}]_j = \frac{1}{N}\sum_{n=1}^{N} r(nT - jT + \phi_0)x(n-1) \tag{20}$$

- The j-th entry of the 16×1 global cross-correlation vector $\mathbf{p}_{zx}$ targeting the symbol $x(n)$ evaluates to:

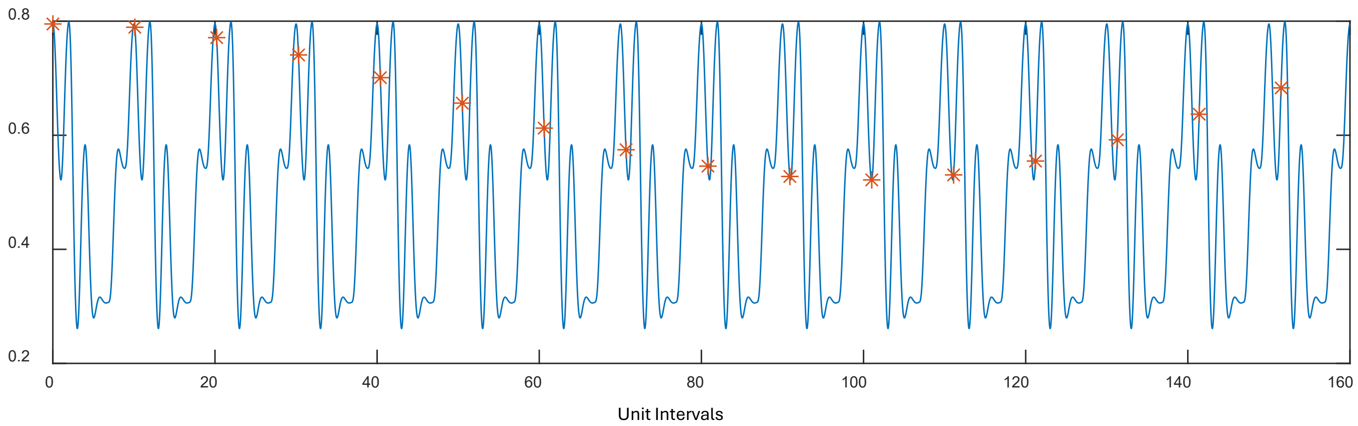
$$[\mathbf{p}_{zx}]_j = \begin{cases} \frac{1}{N}\sum_{n=1}^{N} r(nT - jT + \phi_0)x(n) & \text{for } 0 \leq j \leq 14 \\ \frac{1}{N}\sum_{n=1}^{N} x(n-1)x(n) & \text{for } j = 15 \end{cases} \tag{21}$$

Sampling Scope Operation

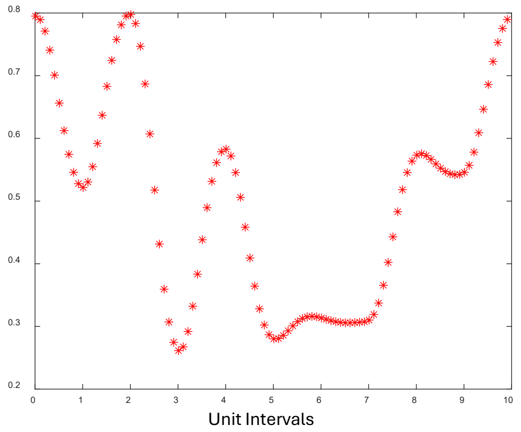
- Sampling scope works with repeating waveform
- Waveform is sampled at a relatively low rate compared to the baud rate of the waveform
- Collects samples at (for example) each repetition of the pattern at a different phase
- Reconstructs waveform by putting samples together
  - Effective sampling rate is determined by spacing of reconstructed samples, much greater than 2 x Nyquist
- Because sampling is not real time, the statistics of the noise change
  - Noise is uncorrelated sample to sample



Example: 10UI PAM4 Waveform



16 Repetitions of Waveform, 1 Sample per Repetition



Reconstructed from 160 Repetitions

Interpolation

- Because the effective sampling rate is so high (compared to Nyquist), interpolation is not a problem to get a sample of the deterministic signal at an arbitrary phase
  - Can use sinc interpolation, spline interpolation, or simple linear interpolation
- However, the noise statistics will be affected by the interpolation
- BUT – the noise statistics have ALREADY been modified by the sampling process

Sampling Scope Effect on Noise

- The reference receiver includes a 4-pole Bessel Thomson filter with 3-dB bandwidth of 1/2T, where T is the symbol period
- The transmitter noise and the scope noise can be represented as white Gaussian noise filtered by the BT filter
- After a few symbol periods, the autocorrelation of the BT filter output goes to zero*
- Since the sampling period of the sampling scope is much greater than a few symbol periods, the noise on the waveform is uncorrelated (white) Gaussian noise
- To accurately model noise on the captured waveform, one must measure the variance of the apparent white noise, remove the white noise, then add back noise of the correct variance filtered by the BT filter

* See contribution: N. Swenson, *Analytical Derivation: Noise Autocorrelation of a 4-Pole Bessel-Thomson Filter*, 2 July 2026

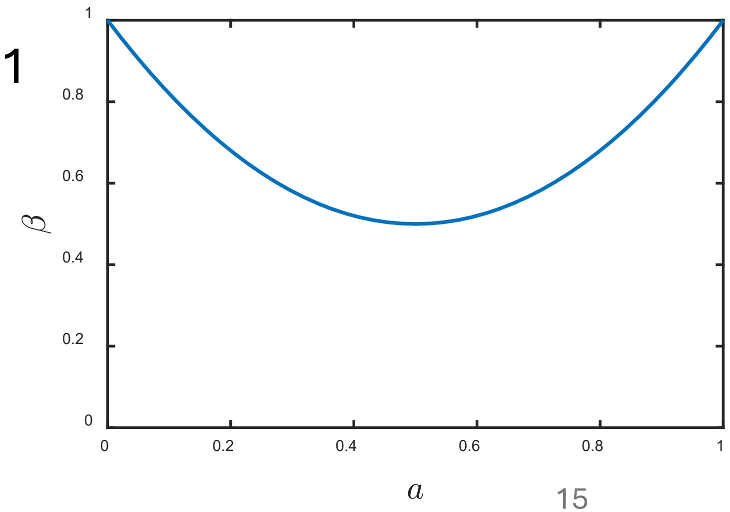
Interpolation Effect on Noise

- As described above, the noise on the waveform captured by the sampling scope appears white (uncorrelated from sample to sample)
- Interpolation is a linear filtering operation that will reduce the noise variance
- Example: linear interpolation
  - y_1+n_1 at time t_1 and y_2+n_2 at time t_2 , where n_1 and n_2 are zero-mean uncorrelated Gaussian noise samples, each of variance σ^2 .
  - Desire sample y_ϕ at time $t_\phi = t_1 + a(t_2 - t_1) = t_1 (1-a) + a\, t_2$, $0 \le a \le 1$

$y_\phi = y_1(1-a) + y_2 a$ $\quad\quad$ $n_\phi = n_1(1-a) + n_2 a$

$$E\left[n_\phi^2\right] = \sigma^2(1-a)^2 + \sigma^2 a^2 = \sigma^2\beta$$

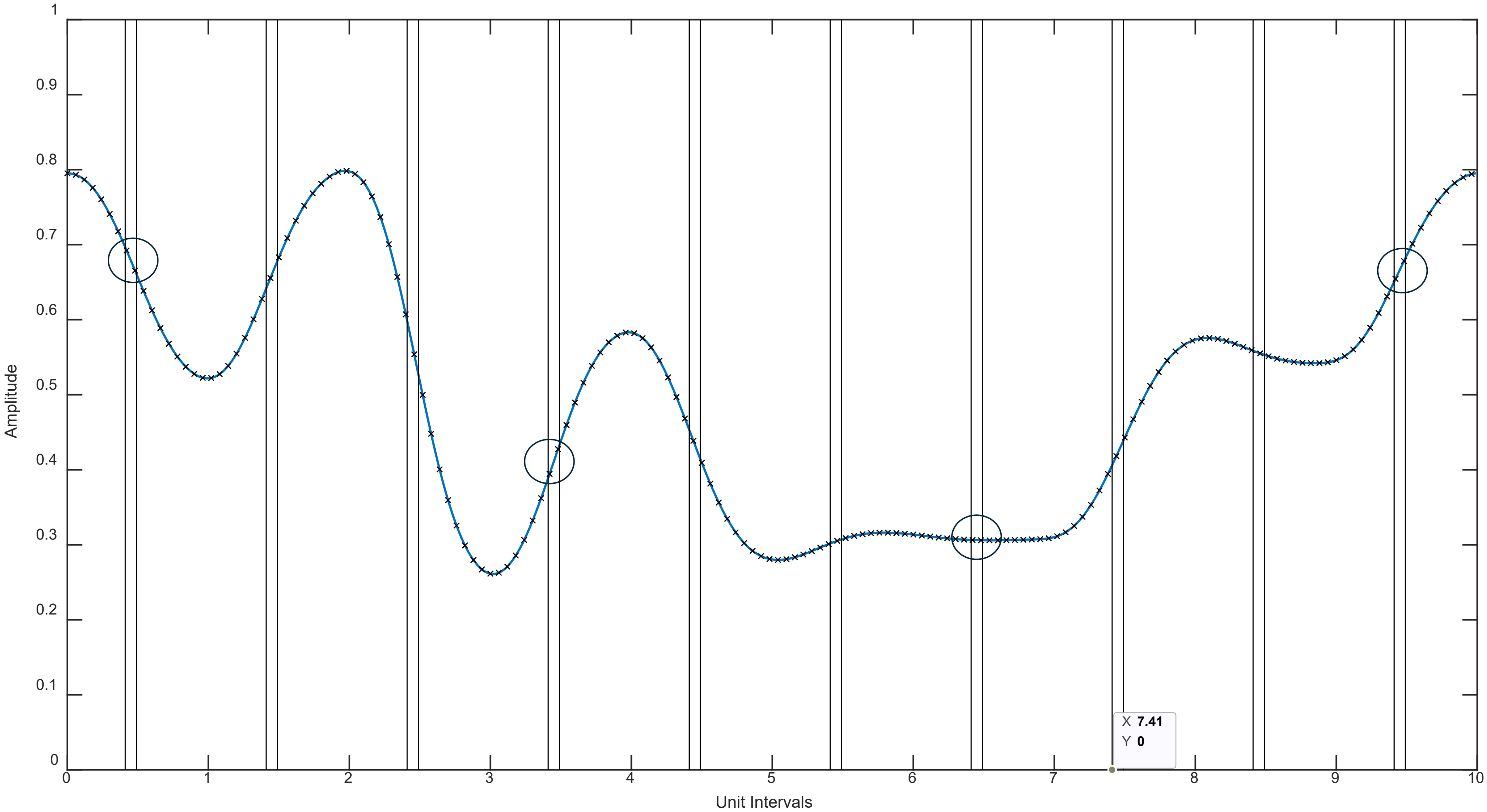
where $\beta = (1 - 2a + 2a^2)$ $\quad\quad$ $1/2 \le \beta \le 1$



Why We Want to Interpolate

- If we don’t interpolate and use finite width histogram bins, we risk missing samples for some symbols or having multiple samples for certain symbols
- Suppose the histogram slice width is T/M. (Currently, M=25.) If the effective sample rate is:
  - Greater than M/T, some slices will capture multiple samples for a given symbol
  - Less than M/T, some slices will have no samples for some symbols
  - Equal to M/T, then there is no need to have a finite width; you get one sample of each symbol in the slice at the same phase

Example: Effective rate > M/T



Summary and Recommendations

- Specify that the histograms are collected at $\phi_0 \pm .05\ T/2$, interpolating if necessary
- Properly account for noise
  - Account for the effects of interpolation, if used
  - De-whiten the noise captured by the sampling scope
  - Details can be added, or we could simply load it into a statement as we do presently:

    "If an equivalent-time sampling oscilloscope is used, the impact of the sampling process and the reference equalizer on transmitter noise has to be compensated for, so that the correct magnitude of noise is present at the output of the equalizer."
  - We could add a sentence:

    "If interpolation is used, the impact of the interpolation process on the statistics of the transmitter noise must also be compensated for so that the correct magnitude of noise is present at the output of the equalizer."

Thank you